MEMORIAL UNIVERSITY OF NEWFOUNDLAND

DEPARTMENT OF MATHEMATICS AND STATISTICS

A4 **MATH 6205** DUE: MARCH 2ND, 2022

Upload your solution to Brightspace at the beginning of class on Monday March 2nd, 2022.

The ultimate goal of this course is for you to be able to read scientific papers on deep learning and being able to implement new methods being proposed in the literature yourself. Here we attempt to do this by implementing *DenseNets* from scratch, and test them for the CIFAR-10 Dataset.

Instructions:

1. Read the original *DenseNet paper*, which can be found here: `https://arxiv.org/pdf/1608.06993.pdf`.
2. There are a few custom building blocks that need to be implemented for a DenseNet model: The *composition function*, the actual *DenseNet block*, and the *transition block*. These can be defined quite straighforwardly using `Keras`' functional API.
3. As DenseNets increase the number of feature maps over each DenseNet block, to make the model more compact, also a compression factor can be implemented for the transition block. Here, use a compression factor of $\theta = 0.5$.
4. Note that in the DenseNet paper, a second version is proposed that also includes *Bottleneck layers*. You can experiment with those as well, but they may not be needed for the relatively small CIFAR-10 Dataset.
5. As actual architecture, use a network that uses a total of 3 DenseNet blocks, with a growth rate of $k = 12$. Process the input images as described in the paper.
6. For the classifier part of the network (i.e. the part after the DenseNet blocks) use global averaging pooling, followed by the dense output layer (as in GoogLeNet).
7. There are two sets of experiments reported in the paper for CIFAR-10. One using data augmentation (as we have seen in class) but no dropout, and one using dropout but no data augmentation. Try out both, using a dropout rate of 20%.
8. Evaluate all your models on the test dataset. Which version gives the most accurate results? Does your DenseNet model beat the other models we have seen in class?
9. **MATH 6205 CIFAR-10 competition (optional!):** Experiment with various hyperparameters (e.g. growth rate, number of layers, number of filters, dropout rate, data augmentation, etc.) and report your best model results by sending me your accuracy score on the testing dataset (train your best model 10 times and send the mean accuracy over all 10 runs). I will keep track of this score and post it on Brightspace.